Part2- Section 3 (Deterministic vs Non Deterministic)

1. Based on the rule-based version, how easy was it to understand how equitably loan decisions are made using Themis?

   ○ Very difficult

   ○ Somewhat difficult

   ○ Neutral

   ○ Easy

   ○ Very Easy

2. Based on the DL-based version, how easy was it to understand how equitably churn decisions are made using Themis?

   ○ Very difficult

   ○ Somewhat difficult

   ○ Neutral

   ○ Easy

   ○ Very Easy

3. For the rule-based version, do Themis generated tests provide additional insights beyond what you can already see in the rules?

   ○ Yes — it confirms discrimination assumptions and reveals corner-cases

   ○ Maybe — rules give transparency, but testing quantifies bias and builds confidence.

   ○ Only when rules are complex — simple rule systems may not benefit as much.

   ○ Absolutely — it uncovers hidden patterns and supports explainability.

   ○ No - It does not help

4. In the DL-based version, do you feel tests generated from Themis are sufficient to explain bias without seeing the internal model training structure?

    ◯ Yes — discrimination tests alone are sufficient to explain model bias.

    ◯ Mostly yes — they provide strong evidence, even without knowing internal training details.

    ◯ Somewhat — they reveal bias patterns, but additional model insights would be useful.

    ◯ Not really — they indicate bias but do not fully explain why the model behaves that way.

    ◯ No — understanding the internal training data and model design is necessary to explain bias.

5. Could Using Themis help you better understand how discrimination can occur in software outcomes?

    ◯ It did not help me understand discrimination at all

    ◯ It slightly helped, but I still found discrimination unclear

    ◯ It somewhat increased my understanding

    ◯ It helped me understand discrimination more clearly

    ◯ It significantly improved my understanding of how discrimination occurs in software outcomes

6. How much control did you feel you had over how Themis conducted and presented discrimination testing?

    ◯ The tool decided everything automatically; I couldn't influence the testing process

    ◯ I could make minor adjustments but had limited influence

    ◯ I could configure certain aspects but not the overall testing process

    ◯ I could meaningfully adjust how tests were conducted and interpreted

    ◯ I had full autonomy to modify and direct the testing and interpretation process

7. How useful do you find defining input subspaces (smaller, focused regions within a large input space) as a method for understanding and analyzing discrimination in complex models?

○ Not useful at all – It provides little or no help in understanding discrimination

○ Slightly useful – It helps in a limited way but lacks clarity or scalability

○ Moderately useful – It provides some insight into discrimination patterns

○ Useful – It meaningfully improves understanding of discrimination across inputs

○ Extremely useful – It is a highly effective and insightful approach for analyzing discrimination in large input spaces